

# FRBRization: αναζητώντας το “κλειδί” για τον εντοπισμό των Έργων

Μανόλης Πεπονάκης, Παναγιώτης Στάικος, Μιχάλης Σφακάκης, Σαράντος Καπιδάκης

*Εργαστήριο Ψηφιακών Βιβλιοθηκών και Ηλεκτρονικής Δημοσίευσης  
Τμήμα Αρχειονομίας – Βιβλιοθηκονομίας, Ιόνιο Πανεπιστήμιο  
{peponakis, pstaikos, msfaka, sarantos}@ionio.gr*

## Περίληψη

Το εννοιολογικό μοντέλο των FRBR αποτελεί ένα πεδίο το οποίο τυγχάνει όλο και μεγαλύτερης προσοχής τόσο μέσα στα περιορισμένα όρια των καταλόγων βιβλιοθηκών όσο και έξω από αυτά ως ένα όχημα που θα επιτρέψει στους καταλόγους των βιβλιοθηκών να αποκτήσουν μια πιο λειτουργική μορφή και να διαλειτουργήσουν και με άλλους φορείς πληροφορίας και περιεχομένου.

Όμως τα εκατομμύρια των βιβλιογραφικών εγγραφών κωδικοποιημένα σε διάφορες διατάξεις είναι εύλογο πως δεν μπορούν να επανακαταλογογραφηθούν. Έτσι αναζητούνται τρόποι μετάβασης των υπάρχουσών εγγραφών στις δομές που ορίζουν οι FRBR. Η παρούσα μελέτη εξετάζει τις διαδικασίες μετάβασης από παραδοσιακές βιβλιογραφικές εγγραφές σε εγγραφές που υιοθετούν τις λογικές των FRBR. Αυτή η διαδικασία, που καλείται FRBRization, διαθέτει πολλές όψεις. Βασικό σημείο εστίασης της εργασίας –και αφετηρία κάθε προσπάθειας FRBRization– αποτελεί η διαδικασία εντοπισμού των βιβλιογραφικών εγγραφών οι οποίες ανήκουν στο ίδιο Έργο.

Εξετάζονται τόσο οι διεθνείς πρακτικές της παραγωγής κλειδιών για τη δημιουργία ομάδων εγγραφών που αποτελούν ένα Έργο, όσο και οι ιδιαίτερες συνθήκες που επικρατούν στους ελληνικούς καταλόγους. Επίσης διερευνάται το πόσο αποτελεσματική θα ήταν η εφαρμογή του αλγορίθμου του OCLC για Work Set Clustering σε ελληνικά δεδομένα. Στη συνέχεια με γνώμονα τόσο τις διεθνείς πρακτικές όσο και τα ελληνικά δεδομένα προτείνονται μερικές βελτιώσεις οι οποίες αυξάνουν την αποτελεσματικότητα της διαδικασίας.

*Λέξεις κλειδιά: FRBR, FRBRization, Κλειδιά ταύτισης*

## Abstract

The conceptual model of FRBR is an area that receives increasing attention, both within the limited boundaries of automated library catalogues and beyond them, as a vehicle that will allow libraries' catalogues to make it towards a more functional interface and interoperate with other organizations who manage information and content.

Unfortunately millions of records -coded in various formats- cannot be re-catalogued. So, it is necessary to develop methods and tools to convert the existing records according to the structures defined by the FRBR. This study examines the procedures for converting traditional bibliographic records into records that have adopted the FRBR logic. This process, namely FRBRization, has many aspects. The focus of our study is the process of detecting bibliographic records belonging to the same *Work*.

The study examines both the international practice of producing keys to create groups of records belonging to the same *Work*, and the specific conditions identified in the Greek on line catalogues. We, also, investigate how effective would be to apply the “OCLC Work Set Clustering algorithm” in the Greek data. Finally, considering both the international practice and the Greek data, we suggest some improvements which increase the efficiency of the overall process.

*Keywords: FRBR, FRBRization, Matching keys*

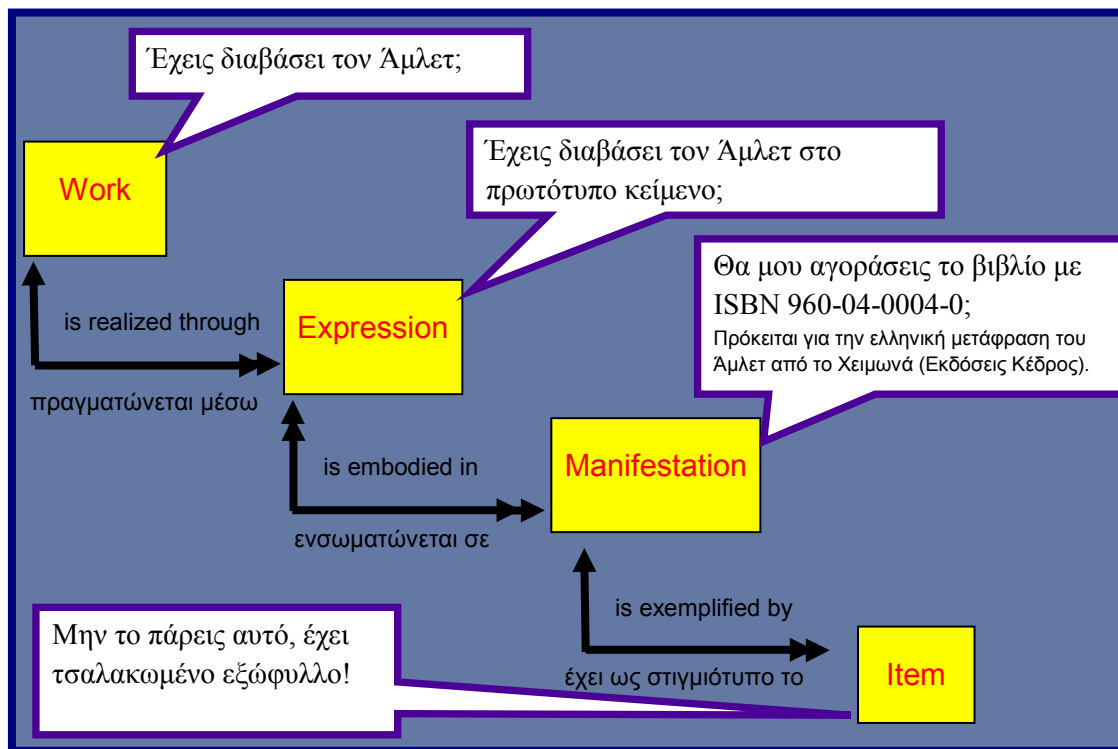
## 1. FRBR

Από το 1992 ως το 1995 η Ομάδα Μελέτης (Study Group) για τις Λειτουργικές Απαιτήσεις για Βιβλιογραφικές Εγγραφές (Functional Requirements for Bibliographic Records) εργάστηκε για την ανάπτυξη ενός μοντέλου σχέσεων οντοτήτων. Το αποτέλεσμα αυτής της εργασίας υπήρξε μια αφαιρετική οπτική του βιβλιογραφικού σύμπαντος ανεξάρτητη από κανόνες καταλογογράφησης ή επιμέρους υλοποιήσεις (Tillet, 2005). Η τελική αναφορά κατατέθηκε το 1997 και εγκρίθηκε από τη “Standing Committee of the IFLA Section on Cataloguing”. Το κείμενο εκδόθηκε το 1998. Οι FRBR, βέβαια, δεν αποτελέσαν μια αποσπασματική προσπάθεια η οποία αποτυπώθηκε σε ένα ορισμένο κείμενο και αφέθηκε στην τύχη της. Η IFLA εξακολουθεί να στηρίζει την προσπάθεια και να την επεκτείνει δημοσιεύοντας το 2009 το [Functional Requirements for Authority Data](#) και εργάζεται πάνω στο [Functional Requirements for Subject Authority Data](#).

Ο πυρήνας του μοντέλου συγκροτείται από: τις οντότητες, τις ιδιότητες που αποδίδονται σε αυτές και τις σχέσεις που έχουν οι οντότητες μεταξύ τους. Το μοντέλο ορίζει δέκα βασικές οντότητες. Για κάθε μία από αυτές ορίζει τις ιδιότητές (attributes) της και τέλος ορίζει τις σχέσεις μεταξύ των οντοτήτων. Οι οντότητες χωρίζονται σε τρεις βασικές κατηγορίες. Η πρώτη κατηγορία περιλαμβάνει τις οπτικές για ένα δημιούργημα και είναι αυτή που

απασχολεί την παρούσα εργασία. Η δεύτερη περιλαμβάνει τους δημιουργούς των οντοτήτων της πρώτης και η τρίτη τη θεματική προσέγγιση.

Μια πιο εύληπτη παρουσίαση για τις οντότητες της πρώτης ομάδας και τις βασικές σχέσεις τους παρουσιάζει με ένα παράδειγμα το σχήμα 1.



Σχήμα 1: Παράδειγμα με τις οντότητες της πρώτης ομάδας και τις σχέσεις τους (Πεπονάκης, 2010: σ.30 )

Οι τέσσερις οντότητες της πρώτης ομάδας εκφράζουν έναν προβληματισμό που υπήρχε για χρόνια στη θεωρία συγκρότησης των καταλόγων. Ήδη από το 1959 η Verona ξεχωρίζει το αντικείμενο από τις εκφράσεις του (αναφέρεται στο O’Neil, 2002). Παρόλα αυτά η βιβλιοθηκονομική κοινότητα εξακολουθούσε να εστιάζει στην περιγραφή των αντικειμένων που είχε στην κατοχή της. Επιχειρώντας μια σύγκριση με τις παραδοσιακές προσεγγίσεις καταλογογράφησης στην ουσία τόσο η έννοια του *Έργου* όσο και της *Έκφρασης* είναι νέες παρότι καταγράφονταν στοιχεία αυτών στο πλαίσιο μιας παραδοσιακής βιβλιογραφικής εγγραφής. Για παράδειγμα ο ομοιόμορφος τίτλος είναι στοιχείο του *Έργου* ενώ η γλώσσα είναι στοιχείο της *Έκφρασης*. Ο βασικός πυρήνας της εγγραφής σχετίζεται με το *Manifestation* ενώ αναφορικά με τα *Αντίτυπα* υπήρχαν τόσο πεδία εντός των βιβλιογραφικών εγγραφών όσο και ολόκληρες εγγραφές οι οποίες συνδέονταν με τις βιβλιογραφικές (π.χ. UNIMARC ή MARC 21 Holdings).

## 2. FRBRization

Είναι ευρέως αποδεκτό πως οι παραδοσιακοί κατάλογοι έχουν φτάσει στα όριά τους και όπως αναφέρει η Yee (2005: σ. 77) είναι σημαντικό να γίνει πιο “έξυπνη” (intelligent) χρήση των εκατομμυρίων εγγραφών που υπάρχουν σε MARC με στόχο να βελτιωθούν οι κατάλογοι

και οι υπηρεσίες που παρέχονται μέσα από αυτούς. Έτσι οι βιβλιοθήκες οφείλουν να αναπτύξουν εργαλεία τα οποία θα είναι αποτελεσματικά μέσα στην ποικιλομορφία των συλλογών και των σχημάτων μεταδεδομένων (Naun, 2010: σ.333). Οι FRBR προσφέρουν μια πιο σύγχρονη αντίληψη ως προς τα βιβλιογραφικά δεδομένα, αλλά όπως αναφέρει ο Rajapatirana (2005) “*re-cataloguing is not an option*”. Κεντρικό ζητούμενο λοιπόν αποτελεί η αξιοποίηση των υπαρχουσών βιβλιογραφικών εγγραφών για την παροχή υπηρεσιών προστιθέμενης αξίας από τις βιβλιοθήκες. Έτσι οδηγούμαστε στην εύρεση μεθόδων οι οποίες θα επιτρέψουν την ανασύνθεση των δεδομένων που υπάρχουν στις νέες μορφές.

Ως FRBRization ορίζεται η διαδικασία ανίχνευσης και σύνθεσης FRBR οντοτήτων χρησιμοποιώντας εγγραφές οι οποίες έχουν καταλογογραφηθεί - κωδικοποιηθεί πρωτογενώς σε άλλα σχήματα κωδικοποίησης. Όπως πολύ εύστοχα παρατηρεί η Babeu (Babeu et al, 2008: σ. 17), οι όροι *FRBR catalog*, *FRBRized system*, *FRBR implementation* χρησιμοποιούνται εναλλακτικά για να περιγράψουν τη διαδικασία αλλά χωρίς ξεκάθαρη σημασία. Η ίδια προτιμά για την περιγραφή της διαδικασίας FRBRization του Perseus να χρησιμοποιεί το περιφραστικό «κατάλογος εμπνευσμένος από τους FRBR».

Αφετηρία της οποιασδήποτε προσπάθειας FRBRization αποτελεί ο εντοπισμός των βιβλιογραφικών εγγραφών οι οποίες αποτελούν ένα *Έργο* και στη συνέχεια ο εντοπισμός εντός αυτής της ομάδας των πιθανών *Εκφράσεων* και *Manifestations*. Για να επιτευχθεί αυτή η ομαδοποίηση παράγονται κάποια «κλειδιά» από τις επιμέρους βιβλιογραφικές εγγραφές και συγκρίνονται μεταξύ τους. Ίδιο κλειδί σημαίνει ίδιο *Έργο*<sup>1</sup>. Με βάση τη διεθνή βιβλιογραφία (Aalberg 2006, Freire 2007, κ.α.) και τα όσα ορίζουν οι FRBR για το τι είναι *Έργο*, τρεις είναι οι βασικές πληροφορίες που πρέπει να περιέχει το κλειδί: ο κύριος δημιουργός του *Έργου*, ο τίτλος του *Έργου* και η μορφή του υλικού (π.χ. κινηματογραφική ταινία ή κείμενο)<sup>2</sup> στην οποία αυτό μπορεί να εκφραστεί.

Με γνώμονα τους FRBR οι τρεις εγγραφές που ακολουθούν συγκροτούν ένα *Έργο* το οποίο αποτελείται από δύο *Εκφράσεις* και τρία *Manifestations*<sup>3</sup>. Με βάση τα όσα έχουν αναφερθεί για τη διαδικασία του FRBRization, το κλειδί που θα συγκέντρωνε τις τρεις εγγραφές κάτω από το ίδιο *Έργο* θα ήταν του τύπου «ΚΕΙΜΕΝΟ – ΟΜΗΡΟΣ 8ΟΣ ΑΙ Π Χ -- ΙΛΙΑΣ».

Εγγραφή 1 - Βιβλίο	
Τίτλος / Συγγραφέας	Ομήρου Ιλιάδα / μετάφραση Ν. Καζαντζάκη, Ι. Θ.Κακριδή
Δημοσίευση	Αθήνα : Εστία, 2000
Ομοιόμορφος τίτλος	Ιλιάς
Συγγραφέας	Όμηρος, 8ος αι. π.Χ.
Μεταφραστής	Καζαντζάκης, Νίκος ; Κακριδής, Ιωάννης Θ.

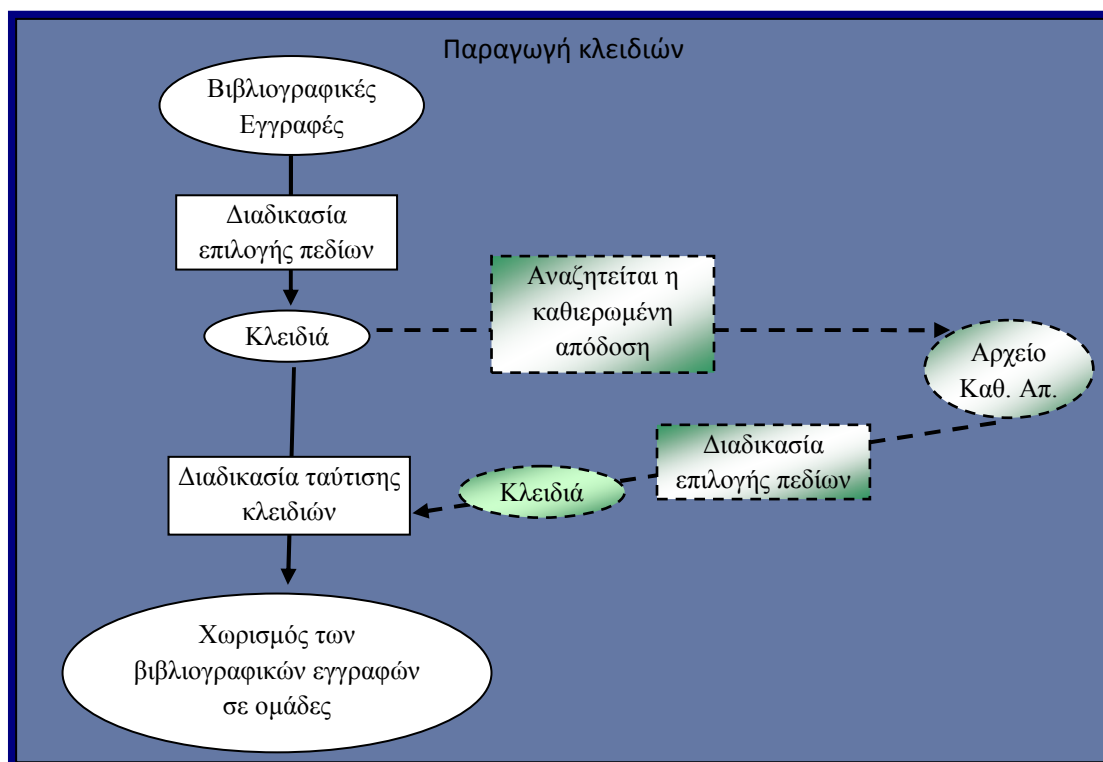
<sup>1</sup> Διαφορετικό κλειδί δεν σημαίνει κατ' ανάγκη και διαφορετικό *Έργο*. Η διαφορά μεταξύ των κλειδιών θα μπορούσε να μετρηθεί με διάφορες μεθόδους ελέγχου ομοιότητας και να ορισθεί κάποιο όριο πάνω από το οποίο οι δύο εγγραφές θα ανήκουν στο ίδιο *Έργο*.

<sup>2</sup> Σε κάποιες περιπτώσεις η μορφή δεν λαμβάνεται υπόψη. Έτσι παράγονται εγγραφές όχι σε επίπεδο Έργου αλλά Σύνολα Έργων (Work Set) όπως τα ονομάζει το OCLC.

<sup>3</sup> Η μία *Εκφραση* είναι η νεοελληνική μετάφραση η οποία περιέχει δύο *Manifestations* (εγγραφές 1 και 2) και η άλλη *Εκφραση* είναι η αγγλική μετάφραση η οποία περιέχει ένα *Manifestation* (εγγραφή 3).

Εγγραφή 2 - Βιβλίο	
Τίτλος / Συγγραφέας	Ιλιάδα / Όμηρος ; έμμετρη μετάφραση Νίκου Καζαντζάκη και Ι.Θ. Κακριδή
Δημοσίευση	Αθήναι : Γερ. Σ. Χρήστου, 1955
Ομοιόμορφος τίτλος	Ιλιάς
Συγγραφέας	Όμηρος, 8ος αι. π.Χ
Μεταφραστής	Καζαντζάκης, Νίκος ; Κακριδής, Ιωάννης Θ.
Εγγραφή 3 - Βιβλίο	
Τίτλος / Συγγραφέας	The Iliad of Homer / translated by Lord Derby
Δημοσίευση	London : Dent, 1944
Ομοιόμορφος τίτλος	Ιλιάς
Συγγραφέας	Όμηρος, 8ος αι. π.Χ
Μεταφραστής	Derby, Lord

Όμως για την παραγωγή των κλειδιών μπορούν να εφαρμοστούν δύο μέθοδοι. Στην πρώτη τα δεδομένα αντλούνται απευθείας από τις βιβλιογραφικές εγγραφές. Στη δεύτερη μεσολαβεί η χρήση Αρχείου Καθιερωμένων Αποδόσεων. Ακολουθεί σχηματική αναπαράσταση και των δύο μεθόδων (με διακεκομμένες γραμμές η χρήση Αρχείου Καθιερωμένων Αποδόσεων).



**Σχήμα 2:** Διαδικασίες δημιουργίας κλειδιών ταύτισης (χωρίς ή με χρήση Αρχείου Καθιερωμένων Αποδόσεων)

Τα πλεονεκτήματα της δεύτερης προσέγγισης για την αποτελεσματικότερη ταύτιση είναι προφανή αφού παρέχεται η δυνατότητα χρήσης επιπλέον πληροφορίας (αυτής του Α.Κ.Α.). Για παράδειγμα είναι δυνατό να ταυτιστεί ο Αριστοφάνης με τον Aristophanes.

### 3. Δημιουργία κλειδιών ταύτισης για εντοπισμό Έργων σε ελληνικές βιβλιογραφικές εγγραφές

Η ενότητα που ακολουθεί έχει ως στόχο να διερευνήσει τη δυνατότητα δημιουργίας κλειδιών τα οποία θα αποτελέσουν τη βάση για την ομαδοποίηση των βιβλιογραφικών εγγραφών που αποτελούν ένα Έργο. Όπως αναφέρθηκε προηγουμένως το κλειδί ταύτισης αποτελείται από τρία βασικά τμήματα. Το πρώτο είναι ο τύπος του υλικού, το δεύτερο ο πνευματικά υπεύθυνος και το τρίτο ο τίτλος. Σχετικά με το πρώτο η τιμή μπορεί να προέρχεται από την ετικέτα εγγραφής η οποία (ως κωδικοποιημένη πληροφορία) δεν επιδέχεται πολλών ερμηνειών ή αμφισημιών. Έτσι η βαρύτητα πέφτει στα άλλα δύο τμήματα του κλειδιού.

Οι συλλογικοί κατάλογοι αποτελούν πολύ καλό πεδίο δοκιμών για FRBRization διότι περιέχουν πολλές *Εκφράσεις* και *Manifestation* ενός Έργου. Στην Ελλάδα υπάρχουν δύο μείζονες συλλογικοί κατάλογοι οι οποίοι αναπτύσσονται με τελείως διαφορετικό τρόπο. Ο ένας είναι ο Συλλογικός Κατάλογος Δημοσίων βιβλιοθηκών και ο άλλος ο Συλλογικός Κατάλογος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών (ΣΚΕΑΒ). Στο Συλλογικό των Δημοσίων η εισαγωγή των δεδομένων ελέγχεται από άνθρωπο και υπάρχει μια πολύ αυξημένη συνέπεια στην καθιέρωση των ονομάτων, από την άλλη όμως υπάρχει σχεδόν παντελής απουσία του πεδίου του ομοιόμορφου τίτλου. Έτσι η φαινομενική ευκολία με την οποία θα λάμβανε χώρα το FRBRization σε έναν συλλογικό κατάλογο ο οποίος έχει δημιουργηθεί με μη αυτόματες διαδικασίες, χάνεται λόγω μη χρήσης του πλέον χρήσιμου για το FRBRization πεδίου: εκείνου του ομοιόμορφου τίτλου. Αντίθετα ο ΣΚΕΑΒ παράγεται αυτόματα χωρίς λεπτομερείς ελέγχους από ανθρώπους σε κάθε εγγραφή. Έχοντας κατά νου τα παραπάνω σε συνδυασμό με το μεγάλο μέγεθος του δεύτερου σε σχέση με τον πρώτο αποφασίστηκε να χρησιμοποιηθεί ο ΣΚΕΑΒ αφού θεωρήθηκε πως είναι μεγαλύτερο και αντιπροσωπευτικότερο δείγμα της ελληνικής κατάστασης.

Η έλλειψη ενός κοινού αρχείου καθιερωμένων αποδόσεων για τα ονόματα είναι γνωστή στις ελληνικές βιβλιοθήκες και έχει τεκμηριωθεί στην ελληνική βιβλιογραφία (βλέπε Άνθη-Καλοφωλιά, 2004). Όμως ακόμη και αν είναι καθιερωμένοι με πολλούς διαφορετικούς τρόπους οι συγγραφείς, τα πεδία των πνευματικά υπεύθυνων είναι συμπληρωμένα στις βιβλιογραφικές εγγραφές. Τι γίνεται όμως με τους ομοιόμορφους τίτλους η ύπαρξη ή όχι των οποίων έχει καταλυτική σημασία για τα Έργα που έχουν πολλές *Εκφράσεις* και *Manifestations* (Hickey and O'Neill, 2005: σ. 246); Μια μικρή δειγματοληψία από τον ΣΚΕΑΒ για μερικά ιδιαίτερα «γνωστά» έργα δίνει τα εξής στοιχεία για τη χρήση του πεδίου του ομοιόμορφου τίτλου.

α/α	Τίτλος / Συγγραφέας	Εγγραφές που αντλήθηκαν από ΣΚΕΑΒ	Εγγραφές με ομοιόμορφο τίτλο	Ποσοστό (%) εγγραφών με ομοιόμορφο τίτλο
1.	Αποκάλυψη / Ιωάννης ο Απόστολος	17	3	17,6
2.	Ερωτόκριτος / Κορνάρος	98	7	7,1
3.	Ευμενίδες / Αισχύλος	94	30	31,9

α/α	Τίτλος / Συγγραφέας	Εγγραφές που αντλήθηκαν από ΣΚΕΑΒ	Εγγραφές με ομοιόμορφο τίτλο	Ποσοστό (%) εγγραφών με ομοιόμορφο τίτλο
4.	Ζορμπάς / Καζαντζάκης	33	11	33,3
5.	Η Πάππισα Ιωάννα / Ροΐδης	70	5	7,1
6.	Θεσμοφοριάζουσες / Αριστοφάνης	43	21	48,9
7.	Ιλιάδα / Όμηρος	40	15	37,5
8.	Καπετάν Μιχάλης / Καζαντζάκης	43	17	40
9.	Ο ζητιάνος / Καρκαβίτσας	45	1	2,2
10.	Ο Χριστός ξανασταυρώνεται / Καζαντζάκης	50	12	24
11.	Οθέλλος / Shakespeare	66	29	43,9
12.	Τελευταίος πειρασμός / Καζαντζάκης	28	7	25

Πίνακας 1: Συχνότητα ύπαρξης ομοιόμορφων τίτλων

Όπως φαίνεται από τον παραπάνω πίνακα η χρήση του ομοιόμορφου τίτλου είναι ιδιαίτερα μικρή ακόμη και για πολύ «διάσημα» Έργα. Για τη χρήση του πεδίου του ομοιόμορφου τίτλου ειδικά στις περιπτώσεις των μεταφράσεων αντιπαρέβαλε με O'Neill (2002).

### 3.1. Δημιουργώντας τα κλειδιά

Όπως αναφέρθηκε ήδη θα πρέπει να δημιουργηθούν κλειδιά τα οποία και θα συγκριθούν μεταξύ τους για να κριθεί κατά πόσο δύο ή περισσότερες εγγραφές ανήκουν στο ίδιο Έργο. Για τη δημιουργία των κλειδιών αναπτύχθηκε μια εφαρμογή λογισμικού η οποία είναι on line και παρέχει τη δυνατότητα στον χρήστη να κάνει upload ένα αρχείο ISO 2709 (UNIMARC ή MARC 21) και το σύστημα να του παράγει τα κλειδιά τα οποία ο χρήστης μπορεί να αποθηκεύσει σε ένα αρχείο κειμένου για περαιτέρω επεξεργασία. Επίσης έχει τη δυνατότητα να διαχειριστεί διαφορετικά σύνολα χαρακτήρων (character sets). Η εφαρμογή όσον αφορά τη σειρά επιλογής πεδίων τίτλου βασίζεται στον αλγόριθμο του OCLC (Hickey and Toves, 2005: σ.5-6)<sup>4</sup> όπου προτείνεται η ακόλουθη σειρά επιλογής:

- Ομοιόμορφος Τίτλος (Κύρια Αναγραφή), (UNIMARC 500, MARC21 130)
- Ομοιόμορφος Τίτλος (Μη κύρια αναγραφή), (UNIMARC 500<sup>5</sup>, MARC21 240)
- Μετάφραση Τίτλου από τον Καταλογογράφο, (UNIMARC 541, MARC21 242)

<sup>4</sup> Πρόκειται για τον αλγόριθμο που χρησιμοποιείται από τα περισσότερα FRBRization projects.

<sup>5</sup> Στο MARC 21 όταν ο ομοιόμορφος τίτλος είναι κύρια αναγραφή καταχωρίζεται στο πεδίο 130 ενώ όταν δεν είναι στο 240. Στο UNIMARC καταχωρίζεται στο ίδιο πεδίο (500) απλά ξεχωρίζει από τον δείκτη 2 (τιμή 0 δεν είναι κύρια αναγραφή, τιμή 1 είναι κύρια αναγραφή).



- Κύριος Τίτλος, (UNIMARC 200, MARC21 245)
- Παραλλαγές του Τίτλου, (UNIMARC 517, MARC21 246)
- Προηγούμενος Τίτλος (UNIMARC 520, MARC21 247)

Για το συγγραφέα επιλέγεται το πεδίο της κύριας αναγραφής και όπου αυτό δεν υπάρχει δημιουργείται ένα από το πρώτο πεδίο εναλλακτικής πνευματικής υπευθυνότητας που έχει κωδικό ρόλου (υποπεδίο \$4) τιμή «070» (Sfakakis and Kapidakis, 2009).

### 3.2. Συγκρίνοντας κλειδιά - Δημιουργώντας τις ομάδες

Για τη διερεύνηση της αποτελεσματικότητας της διαδικασίας εντοπισμού *Έργων* επιλέχθηκαν τρία ευρέως γνωστά *Έργα*: ένα της νεοελληνικής λογοτεχνίας και δύο της αρχαίας ελληνικής γραμματείας. Πρόκειται για την Πάπισσα Ιωάννα του Ροΐδη, την Ιλιάδα του Ομήρου και τις Θεσμοφοριάζουσες του Αριστοφάνη. Συνολικά και για τα τρία *Έργα* αντλήθηκαν από τον ΣΚΕΑΒ 153 βιβλιογραφικές εγγραφές.

Για την Πάπισσα Ιωάννα του Ροΐδη εντοπίστηκαν 70 εγγραφές. Οι εγγραφές αυτές χωρίζονται σε 62 στην ελληνική γλώσσα και 8 μεταφράσεις σε 4 γλώσσες (3 αγγλικές, 3 γερμανικές, 1 γαλλική και 1 ιταλική). Δηλαδή με όρους FRBR προκύπτουν: 1 *Έργο*, 5 Εκφράσεις (τουλάχιστον)<sup>6</sup> και 70 *Manifestations* (με την προϋπόθεση ότι το de-duplication του ΣΚΕΑΒ είναι σωστό)

Επειδή ο όγκος των δεδομένων για την Ιλιάδα του Ομήρου είναι τέτοιος (περισσότερες από 1000 εγγραφές) που το σύστημα διαχείρισης του ΣΚΕΑΒ δεν επέτρεπε την αποθήκευσή τους για περαιτέρω επεξεργασία επιλέχθηκε να γίνει δειγματοληψία<sup>7</sup>. Έτσι αντλήθηκε δείγμα 40 εγγραφών της Ιλιάδας του Ομήρου και εντοπίστηκαν οι εξής γλώσσες: 3 στην αγγλική, 3 στην αγγλική με νέα ελληνικά, 3 στην αγγλική με αρχαία ελληνική, 1 στη γαλλική, 2 γαλλική με αρχαία ελληνική, 6 αρχαία ελληνική, 3 γερμανική με αρχαία ελληνική, 3 αρχαία ελληνική με λατινική, 10 νέα ελληνική, 4 νέα και αρχαία ελληνική, 1 τουρκική με αρχαία ελληνική και 1 χωρίς δήλωση γλώσσας.

Για το *Έργο* Θεσμοφοριάζουσες του Αριστοφάνη εντοπίστηκαν συνολικά 43 εγγραφές. Μέσα στις 43 εγγραφές εντοπίστηκαν οι εξής γλώσσες: 4 στην αγγλική, 10 στην αγγλική με αρχαία ελληνική, 3 αρχαία ελληνική, 4 νέα ελληνική, 17 αρχαία ελληνική με νέα ελληνική, 1 ιταλική, 3 αρχαία ελληνική με ιταλική και 1 χωρίς δήλωση γλώσσας.

#### 3.2.1 Ονόματα Συγγραφέων

Στα έργα της Πάπισσας Ιωάννας το όνομα του Ροΐδη εμφανίζεται -μετά από κανονικοποίηση- με τέσσερις διαφορετικούς τρόπους ως εξής:

<sup>6</sup> Με δεδομένες τις δυσκολίες εντοπισμού των Εκφράσεων ακολουθείται η απλουστευμένη λογική της ομαδοποίησης ανά γλώσσα και μορφή.

<sup>7</sup> Για τη μέθοδο δειγματοληψίας με πίνακες τυχαίων αριθμών βλέπε «Ρίτσαρντσον, Σ. Κ. και Βασιλείνας, Α. (2009). Εισαγωγή στη στατιστική επιστήμη. Αθήνα: Κάκτος : σ. 217-222»



α/α	Όνομα μετά από κανονικοποίηση	Εγγραφές που αντιστοιχούν
1.	ΡΟΙΔΗΣ ΕΜΜΑΝΟΥΗΛ 18361904	23
2.	ΡΟΙΔΗΣ ΕΜΜΑΝΟΥΗΛ	3
3.	ΡΟΙΔΗΣ ΕΜΜΑΝΟΥΗΛ Δ 18361904	17
4.	ΡΟ ΙΔΗΣ ΕΜΜΑΝΟΥΗΛ Δ 18361904	27

**Πίνακας 2:** Παραλλαγές του ονόματος του Ροΐδη

Όπως φαίνεται στην 4<sup>η</sup> περίπτωση τα διαλυτικά κυριολεκτικά διέλυσαν το όνομα!

Στις εγγραφές του δείγματος το όνομα του Ομήρου εμφανίζεται, μετά από κανονικοποίηση, όπως φαίνεται στον παρακάτω πίνακα.

α/α	Όνομα Συγγραφέα μετά από κανονικοποίηση	Εγγραφές που αντιστοιχούν
1.	HOMER	1
2.	ΒΟΜΗΡΟΣ 8ΟΣ ΑΙ ΠΧ	2
3.	ΟΜΗΡΟΣ	1
4.	ΟΜΗΡΟΣ 8ΟΣ ΑΙ ΠΧ	20
5.	ΟΜΗΡΟΣ 8ΟΣ ΑΙΩΝΑΣ ΠΧ	12
6.	ΟΜΗΡΟΣ ΠΕΡ 8ΟΣ ΑΙ ΠΧ	4

**Πίνακας 3:** Παραλλαγές του ονόματος του Ομήρου

Στις εγγραφές του δείγματος το όνομα του Αριστοφάνη εμφανίζεται με 9 διαφορετικούς τρόπους ως εξής:

α/α	Όνομα Συγγραφέα μετά από κανονικοποίηση	Εγγραφές που αντιστοιχούν
1.	ARISTOPHANES	6
2.	ARISTOPHANES CA 446CA 385 BC	5
3.	ΑΡΙΣΤΟΦΑΝΗΣ 438338 ΠΧ	3
4.	ΑΡΙΣΤΟΦΑΝΗΣ 444388	4
5.	ΑΡΙΣΤΟΦΑΝΗΣ 444388 ΠΧ	12
6.	ΑΡΙΣΤΟΦΑΝΗΣ 446385 ΠΧ	9
7.	ΑΡΙΣΤΟΦΑΝΗΣ 450380 ΠΧ	1
8.	ΑΡΙΣΤΟΦΑΝΗΣ 450385 ΠΧ	1
9.	ΑΡΙΣΤΟΦΑΝΗΣ Π 446385 ΠΧ	2

**Πίνακας 4:** Παραλλαγές του ονόματος του Αριστοφάνη

Όπως προκύπτει από τους παραπάνω πίνακες η συνέπεια ως προς τη γραφή του ονόματος είναι αρκετά μειωμένη αφού το όνομα κάθε συγγραφέα εμφανίζεται στις βιβλιογραφικές εγγραφές του δείγματος με 6,3 διαφορετικούς τρόπους κατά μέσο όρο.

### 3.2.2 Τίτλοι

Ο κύριος τίτλος κάθε έργου (πεδίο 200 UNIMARC) υπάρχει στην 4<sup>η</sup> θέση του αλγορίθμου του OCLC (βλέπε παραπάνω) ενώ την πρώτη θέση κατέχει ο ομοιόμορφος τίτλος, η χρήση του οποίου είναι αρκετά περιορισμένη όπως προκύπτει από τον πίνακα 1.

Στις 153 εγγραφές του δείγματος των τριών *Έργων* υπάρχουν ομοιόμορφοι τίτλοι σε 41 εγγραφές (ποσοστό 26,7). Μια πιο προσεκτική ματιά στα δεδομένα θα δείξει πως ακόμη και η περιορισμένη παρουσία πεδίων ομοιόμορφου τίτλου ομογενοποιεί τα δεδομένα σε βαθμό ακόμη μικρότερο από τον αναμενόμενο. Στην περίπτωση της *Ιλιάδας* ο ομοιόμορφος τίτλος δεν ήταν και πολύ “ομοιόμορφος”. Στις 15 εγγραφές που υπήρχε είχε τις εξής 4 μορφές: *Ιλιάδα*, *Ιλιάς*, *Iliad*, *Ilias*. Το ίδιο πρόβλημα εμφανίζεται και στην περίπτωση του *Αριστοφάνη* όπου ο ομοιόμορφος τίτλος εμφανίζεται με τις εξής 3 μορφές: *Thesmophoriazusae*, *Thesmophoriazousai*, *Θεσμοφοριάζουσαι*. Στην περίπτωση της *Πάππιας* *Ιωάννας* του *Ροΐδη* τα πράγματα είναι σαφώς καλύτερα αφού οι δύο μόλις παραλλαγές με τις οποίες εμφανίζεται ο ομοιόμορφος τίτλος διαφέρουν μεταξύ τους στην παρουσία ή απουσία του άρθρου «Η». Έτσι προκρίνοντας τον ομοιόμορφο τίτλο (όπου αυτός υπάρχει) σε σχέση με τον κύριο προκύπτουν για το *Έργο* *Πάππιας* *Ιωάννα* 15 διαφορετικά κλειδιά, για την *Ιλιάδα* 23 και για τις *Θεσμοφοριάζουσες* 6.

Σε μια προσπάθεια βελτίωσης του αλγορίθμου του OCLC και προσαρμογής του στην ελληνική πραγματικότητα επιχειρήθηκε η προσθήκη στη σειρά επιλογής του αλγορίθμου του πεδίου διασύνδεσης “Μετάφραση τού” (454 UNIMARC). Οι αριθμοί δείχνουν μια σαφή (αλλά όχι θεαματική) βελτίωση των αποτελεσμάτων. Συγκεκριμένα καταφέρνουν να μειώσουν τα κλειδιά από 15 σε 12 για την *Πάππιας*, από 23 σε 22 στην *Ιλιάδα*, ενώ αφήνουν σταθερά στο 6 αυτά από τις *Θεσμοφοριάζουσες*.

### 3.2.3 Τίτλοι και ονόματα μαζί - κλειδιά ανά περίπτωση

Σε έναν βιβλιοθηκονομικά ιδεατό κόσμο τα κλειδιά για κάθε *Έργο* θα ταυτίζονταν μεταξύ τους ώστε να μένει τελικά ένα κλειδί για την *Πάππιας* *Ιωάννα*, ένα για την *Ιλιάδα* και ένα για τις *Θεσμοφοριάζουσες*. Αυτό θα συνέβαινε διότι πολύ απλά το όνομα του συγγραφέα θα ήταν γραμμένο με τον ίδιο ακριβώς τρόπο σε όλες τις εγγραφές του συγγραφέα και ο ομοιόμορφος τίτλος επίσης θα ήταν γραμμένος ομοιόμορφα και θα υπήρχε. Βέβαια τέτοια ποσοστά επιτυχίας σπανίως καταγράφονται από τη διεθνή βιβλιογραφία όταν υπάρχουν μεγάλοι όγκοι δεδομένων. Όμως η προσπάθεια ταιριάσματος (matching) στα δεδομένα του ΣΚΕΑΒ δίνει αποτελέσματα που τείνουν προς την αντίθετη κατεύθυνση. Όπως φαίνεται και στον παρακάτω πίνακα αντί για ένα κλειδί στην *Ιλιάδα* του *Ομήρου* παράγονται 28 διαφορετικά, στις *Θεσμοφοριάζουσες* παράγονται 17 και στην *Πάππιας* *Ιωάννα* 22.

	Πάππιας Ιωάννα	Ιλιάδα	Θεσμοφοριάζουσες
Διαφορετικά κλειδιά / Εγγραφές δείγματος	<b>22/70 = 0,31</b> Περίπου κάθε 3 εγγραφές και άλλο κλειδί	<b>28/40 = 0,7</b> Ούτε δύο όμοια κλειδιά ανά τρεις εγγραφές	<b>17/43 = 0,39</b> Κάθε 5 εγγραφές και δύο κλειδιά

Πίνακας 5: Σχέση εγγραφών ως προς τα κλειδιά που παράγονται

Είναι βέβαια γεγονός πως ο αριθμός των κλειδιών είναι αντιστρόφως ανάλογος του αριθμού των εγγραφών. Αυτό συμβαίνει διότι περιορίζεται η “πολυφωνία” των ομοιόμορφων τίτλων και ονομάτων. Πάντως αξίζει να σημειωθεί πως αν ακολουθούνταν με συνέπεια αρχεία καθιερωμένων αποδόσεων, παρά τη σχετικά μικρή παρουσία ομοιόμορφων τίτλων, τα αποτελέσματα θα ήταν πολύ διαφορετικά. Συγκεκριμένα τα κλειδιά θα διαμορφώνονταν ως εξής: τα 22 διαφορετικά κλειδιά της Πάππισας Ιωάννας θα γίνονταν 12, τα 28 διαφορετικά κλειδιά της Ιλιάδας θα γίνονταν 20 και τα 17 από τις Θεσμοφοριάζουσες θα γίνονταν μόλις 5. Δηλαδή τα ποσοστά επιτυχίας σχεδόν θα διπλασιάζονταν.

#### 4. Συμπεράσματα

Σε βιβλιογραφικά δεδομένα τα οποία θα συμμορφώνονταν απολύτως με τα όσα επιτάσσουν οι βιβλιοθηκονομικοί κανόνες γύρω από τη δημιουργία και χρήση των σημείων πρόσβασης ο εντοπισμός των *Έργων* θα αποτελούσε μια σχετικά εύκολη διαδικασία κυρίως για τα *Έργα* εκείνα των οποίων οι πολλές εκδόσεις επιτάσσουν τη δημιουργία και χρήση ομοιόμορφων τίτλων. Η ελληνική πραγματικότητα πολύ απέχει από μια τέτοια κατάσταση. Η χρήση των κατ’ όνομα ομοιόμορφων τίτλων είναι αποσπασματική και οι τίτλοι κάθε άλλο παρά ομοιόμορφοι είναι. Άρα η προσέγγιση του απλού ταιριάσματος (matching) φαίνεται να οδηγεί σε πενιχρά αποτελέσματα. Η βελτίωση που επιτυγχάνεται με την προσθήκη του πεδίου «Μετάφραση τού» δεν επαρκεί.

Έτσι, επιβεβαιώνεται για άλλη μια φορά τόσο η ανυπαρξία χρήσης αρχείου καθιερωμένων αποδόσεων κατά την καταλογογράφηση, όσο και η αναγκαιότητα χρήσης του. Μένει να διερευνηθεί κατά πόσο θα βελτίωνε τα αποτελέσματα η εκ των υστέρων χρήση ενός αρχείου καθιερωμένων αποδόσεων κατά την παραγωγή των κλειδιών (βλέπε σχήμα 2). Βέβαια τα ποσοστά βελτίωσης θα ήταν άρρηκτα συνδεδεμένα με την ποιότητα του αρχείου καθιερωμένων αποδόσεων.

Ένας ακόμη τρόπος για τη βελτίωση των αποτελεσμάτων είναι η χρήση τεχνικών ελέγχου βαθμού ομοιότητας (similarity) αν και πάλι προβληματίζουν οι μεγάλες αποκλίσεις στους τρόπους απόδοσης ονομάτων και ομοιόμορφων τίτλων. Ένα από τα βασικά προβλήματα είναι οι διαφορετικές γλώσσες. Όπως φαίνεται και στον πίνακα 4 κάτω από την αγγλόφωνη γραφή του ονόματος του Αριστοφάνη εμφανίζονται οι 11 από τις 43 εγγραφές. Επίσης κάτω από ξενόγλωσσους ομοιόμορφους τίτλους εμφανίζονται 5 (από τις 40) βιβλιογραφικές εγγραφές στην Ιλιάδα και 8 (από τις 43) στις Θεσμοφοριάζουσες. Με τους κλασικούς αλγορίθμους μέτρησης ομοιότητας οι εν λόγω εγγραφές δεν είναι δυνατό να ταυτιστούν με αυτές της ελληνικής γραφής.

#### Βιβλιογραφία

Aalberg, T. (2006). A Process and Tool for the Conversion of MARC Records to a Normalized FRBR Implementation. In: ICADL 2006, Kyoto, Japan, 27-30 November 2006. Sugimoto, S. et al. (eds.) Berlin, Heidelberg: Springer, pp. 283-292.

Babeu, A. (2008). Building a "FRBR-Inspired" Catalog: The Perseus Digital Library Experience. [Internet] Perseus Digital Library. Διαθέσιμο στο <http://www.perseus.tufts.edu/~ababeu/PerseusFRBRExperiment.pdf>.

Freire, N., Borbinha, J. and Calado, P. (2007). Identification of FRBR Works Within Bibliographic Databases: An Experiment with UNIMARC and Duplicate Detection Techniques. **In: ICADL 2007, Hanoi, Vietnam, 10-13 December 2007**. Berlin, Heidelberg: Springer, pp. 267–276.

Hickey, T.B. and O'Neill, E.T. (2005). FRBRizing OCLC's WorldCat. Cataloging & Classification Quarterly. **39** (3/4), pp. 239-251.

Hickey, T.B. and Toves, J. (2005). FRBR Work-Set Algorithm. Διαθέσιμο στο [http://www.oclc.org/research/software/frbr/frbr\\_workset\\_algorithm.pdf](http://www.oclc.org/research/software/frbr/frbr_workset_algorithm.pdf) (απαιτείται εγγραφή).

Hickey, T.B., O'Neill, E.T and Toves, J. (2002). Experiments with the IFLA Functional Requirements for Bibliographic Records (FRBR). D-Lib Magazine. 8, (9).

IFLA Study Group on the Functional Requirements for Bibliographic Records. (1998). Functional requirements for bibliographic records: final report. IFLA. Διαθέσιμο στο <http://www.ifla.org/files/cataloguing/frbr/frbr.pdf>.

Naun, C. C. (2010). Next Generation OPACs: A Cataloging Viewpoint. Cataloging & Classification Quarterly, 48 (4), pp. 330 - 342

O'Neill, E.T. (2002). FRBR: Functional Requirements for Bibliographic Records Application of the Entity-Relationship Model to Humphry Clunker. Library Resources & Technical Services. 46 (4), pp. 150-159.

Rajapatirana, B. and Missingham, R. (2005). The Australian National Bibliographic Database and the Functional Requirements for the Bibliographic Database (FRBR). The Australian Library Journal. 54 (1), pp. 31-42.

Sfakakis, M. and Kapidakis, S. (2009). Eliminating query failures in a work-centric library meta-search environment. Library Hi Tech. 27 (2), pp. 286-307.

Tillett, B. (2005). What is FRBR? A conceptual model for the bibliographic universe. The Australian Library Journal. Διαθέσιμο στο <http://www.alia.org.au/publishing/alj/54.1/full.text/tillett.html>.

Yee, M.M. (2005). FRBRization: a Method for Turning Online Public Finding Lists into Online Public Catalogs. Information Technology and Libraries. 24 (3), pp. 77-95.

Άνθη – Καλοφωλιά, Έ. (2004). Αξιοπιστία και αποτελεσματικότητα των ελληνικών καταλόγων: ασυνέπειες και λάθη στα αρχεία καθιερωμένων ονομάτων. **Στο: 13<sup>ο</sup> Πανελλήνιο Συνέδριο Ακαδημαϊκών Βιβλιοθηκών. Κέρκυρα, 13-15 Οκτωβρίου 2004**. Διαθέσιμο στο <http://www.ionio.gr/libconf/pdfs/Anthi.pdf>.

Πεπονάκης, Μ. (2010). Σύνθεση FRBR εγγραφών αξιοποιώντας υπάρχουσες βιβλιογραφικές εγγραφές (FRBRization): ομαδοποίηση σχετικών εγγραφών (clustering) και εμφάνισή τους σε on line συστήματα. Μεταπτυχιακή Εργασία, Ιόνιο Πανεπιστήμιο - Τμήμα Αρχιτεκτονικής Βιβλιοθηκονομίας. Διαθέσιμο στο [http://dlib.ionio.gr/mtheses/Peponakis\\_FRBRization.pdf](http://dlib.ionio.gr/mtheses/Peponakis_FRBRization.pdf)